

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

APPLICATION FOR U.S. LETTERS PATENT

Title:

**TRANSLATION LOOK ASIDE BUFFER (TLB) WITH INCREASED
TRANSLATIONAL CAPACITY FOR MULTI-THREADED COMPUTER
PROCESSES**

Inventors:

Jeffrey Todd Bridges
Les M. DeBruyne
Robert L. Goldiez
Michael S. McIlvaine
Thomas A. Sartorius
Rodney Wayne Smith

George R. Pettit - 27,369
CONNOLLY BOVE LODGE & HUTZ LLP
1990 M Street, N.W., Suite 800
Washington, DC 20036-3425
(202) 331-7111

**TRANSLATION LOOK ASIDE BUFFER (TLB) WITH INCREASED
TRANSLATIONAL CAPACITY FOR MULTI-THREADED COMPUTER
PROCESSES**

BACKGROUND OF THE INVENTION

[0001] The present invention relates to digital computer systems which provide multi-threaded execution. Specifically, a translation look-aside buffer (TLB) is provided which reduces the number of entries in the main memory required to service a multi-threaded computer system.

[0002] In order to increase the overall speed of computer program execution, multi-threaded computer processing units execute a plurality of threads associated with the program at one time. The execution of the program is divided into multiple threads which are active at the same time, and various hardware resources of the processor can simultaneously execute the active threads. Simultaneous processing of multiple independent instruction streams keeps the processors computational hardware resources active. Improved execution efficiency results and normal pipe-line stalls that may occur with a single threaded processor which are precipitated by instruction dependencies can be avoided with multi-threaded computer processors.

[0003] High performance multi-threaded processors have instructions from multiple threads which are in progress at the same time in different parts of the execution pipe line. Each of the threads is identified as a context and is allocated physical storage elements to hold the state associated with the thread. In any one instance, there is a physical register to hold an executing thread's architectural context. In this way, the various processes being executed are tagged with a thread ID, so that the computing results associated with each thread context can be applied to the correct architectural resources in the multi-threaded system.

[0004] In both single threaded processors and multi-threaded processors, memory management is necessary so that the program can retrieve values stored in a memory relatively quickly. A common technique used in memory management employs a look-aside buffer (TLB) which caches address translation key pairs. The TLB is generally a content addressable memory (CAM) having a virtual address as its look-up key. Program execution identifies a virtual address which is translated by the look-aside buffer (TLB) to obtain a real address of a memory location of a value needed for the program thread execution.

[0005] Entries in the translational look-aside buffer (TLB) are generally organized so that a virtual page number identified from code execution identifies a real page number stored within the memory. The TLB identifies from a virtual page number a group of pages, starting at a location identified by the virtual page number (VPN). The location within the group of pages is identified by the lower order bits of the virtual page number to save space in the look-aside buffer (TLB). This is important, in that the translational look-aside buffer (TLB) is a hardware table with a fixed capacity and if the CPU uses more pages of memory than the

number of TLB mapping cache entries, the TLB will have to be updated from an external memory. The process of accessing the external memory and obtaining updates slows down the memory management process, and thus the overall relative speed of execution. With many threads running on the CPU nearly simultaneously, each of the active threads must keep a set of active mappings in the translational look-aside buffer (TLB) to avoid any significant penalty from fetching the mappings that are not resident in the TLB. Unfortunately, increasing the number of entries in the translational look-aside buffer (TLB) increases the required chip area and increases the access time and power consumption of the translational look-aside buffer (TLB).

[0006] It is therefore desirable to organize the contents of the translational look-up buffer (TLB) to reduce the need for frequent updates of the stored information without increasing the total number of memory locations available for translational data.

SUMMARY OF THE INVENTION

[0007] A method and apparatus for increasing the number of real memory addresses accessible through a translation look-aside buffer (TLB) is provided. Each entry in the TLB includes a virtual address, a real address of a memory location and a special thread implicit mode bit to indicate whether the virtual address represents one of a plurality of threads being processed. When a virtual address in the buffer corresponds to a virtual address sought during processing by the CPU, the real address is read from the buffer entry corresponding to the virtual address. When the special mode bit is set to indicate that one of a plurality of

threads is being processed by the CPU, the CPU concatenates with the higher order bits of the real address a value representing a thread being processed. In the event that the special mode bit is not set, meaning that the buffer entry represents a conventional translation look-aside buffer entry, the entire real address including its lower order bits is used to identify a memory storage area to acquire data for the processor. The real address may be further concatenated with the lower order bits of the virtual address to provide additional granularity to the real address.

[0008] The invention is particularly useful in multi-thread CPU processing. By using the thread identification as part of the real address, a single translation look-aside buffer (TLB) entry can be used to identify multiple addresses corresponding to the number of threads being processed. The invention can co-exist simultaneously with conventional translational look-aside buffer entries by setting the thread implicit mode bit to zero. When this happens, virtual addresses are mapped to real addresses which are unique to a single virtual address.

DESCRIPTION OF THE FIGURES

[0009] Fig. 1 shows a multi-thread processor (CPU) and a memory management unit employing a look-aside buffer (TLB);

[0010] Fig. 2 represents the components of a look-aside buffer (TLB);

[0011] Fig. 3 represents the decoding of a virtual address into a real address in accordance with a preferred embodiment of the invention; and

[0012] Fig. 4 illustrates in flowchart form the process executed by the translational look-ahead buffer (TLB) in accordance with a preferred embodiment.

DESCRIPTION OF THE PREFERRED EMBODIMENT

[0013] Referring now to Fig. 1, a block diagram representation of a multi-thread processor is shown with an accompanying translational look-up buffer (TLB) in accordance with a preferred embodiment of the invention. The multi-thread CPU computing system 10 is illustrated as multiple processors 11A-11N, representing N threads of a multi-thread CPU. Each of the threads 11A-11N execute in a pipeline processor. During simultaneous execution of the various threads, access to a main memory 17 may be required to complete executing an instruction.

[0014] The access to the main memory 17 is through a memory management unit 13. Associated with the memory management unit 13 is a translational look-aside buffer (TLB) 15. The process of retrieving and writing data to the memory 17 is aided with the memory management unit 13 translational look-up buffer (TLB) 15. There is one instance of physical registers to hold each of the executing threads architectural contexts. For each thread, that may be simultaneously executed, there is a copy of the GPRS, the LR, the CTR, the XER and the CR. Each thread being processed is allocated some dedicated physical storage elements in the main memory 17 to hold the state associated with the thread represented by

the contents of the architecturally defined registers. Each instruction that is in the CPU pipeline is tagged with a thread ID so that the architectural results that it produces can be applied to the correct thread's architectural resources. Thread ID register 12 maintains the thread identification ID so that the result of execution can be identified with a particular thread.

[0015] The memory management unit 13 operates from addresses which are visible to the programmer, referred to as the effective address (EA). The effective address (EA) is related to the real address (RA) in main memory through a translational look-aside buffer (TLB) 15. Multiple instances of a program may be running in a time slice manner, and each program instance can appear to the programmer to use the same memory addresses, but refer to different physical storage locations using the mapping of translational look-aside buffer (TLB) 15.

[0016] In one implementation of a translational look-aside buffer 15, the effective address (EA) is used with a process identifier (PID) which is unique for each process instance. A virtual address (VA) is formed by the concatenation of the effective address (EA) and the process identifier (PID). Together, these entities constitute a one-to-one mapping between a virtual address and a corresponding real address in main memory 17.

[0017] The organization of the translational look-aside buffer (TLB) 15 in accordance with a preferred embodiment is illustrated more particularly with respect to Fig. 2. The process identifier (PID) and effective address (EA) are stored as a virtual address along with a real address for main memory 17 in a

content addressable memory 22. The virtual address comprising process identifier (PID) and the effective address (EA) are used as the look-up keys. The translational buffer 22 includes a search engine 25 which uses a portion of the virtual address comprising the process identifier (PID) and the higher order bits of the effective address (EA) to locate a particular real address (RA) stored in a location 23. Fig. 2, unlike the conventional translational look-up buffer architecture, includes a thread, implicit identifier bit (Tbit). The Tbit is used to identify whether or not the entry 23 in the TLB is associated with one of the multiple threads 11A-11N being executed.

[0018] The search engine 25 may be employed to derive a real address for the main memory 17 using a virtual page number (VPN) derived from the virtual address (VA) and the thread ID identifying the thread for which access to the main memory is being made. Fig. 3 illustrates how the translational look-aside buffer (TLB) entry 23 can be decoded to represent a plurality of real addresses for memory 17.

[0019] The virtual address (VA) which is called for by the executing thread has bits 0-39 which comprise (as shown in Fig. 3) the process identification number (PID 0:7) bits the higher order bits of the effective address (EA 0:21) bits, and the lower order (EA 22:31) bits. In accordance with the present embodiment of the invention, the virtual page number (VPN) comprises the process identification number (PID 0:7) and the higher order bits of the effective address (EA 0:21). Using the VPN (0:29), the RPN, representing the higher order bits of the real address RA (0:22) are located in the memory location 23.

[0020] The remaining portion of the real address of a storage location in main memory 17 is either the lower order bits of the real address RPN (22:25) or the thread ID. As shown in Fig. 3, gate 34 under control of the value of the thread implicit mode bit (Tbit) selects either the thread ID from register 12 or the lower order bits of the real address RPN (22:25). A second logic circuit in the buffer hardware 35 concatenates the result obtained from gate 34 with the lower order bits of the effective address EA (22:31). When the thread implicit mode bit (Tbit) is set to one, the real address for the location in main memory 17 comprises RPN (0:22), the thread ID, and the lower order bits of the effective address EA (22:31). In the event that the thread implicit mode bit has not been set, the gate 34 inserts the real address lower order bits of the real address RPN (22:25) instead of the thread ID (0:2 in the case of an 8 bit thread ID).

[0021] Thus, the same virtual address may be used to identify a group of pages, wherein the particular page within the group is identified by either the lower order bits of the real address or the thread ID.

[0022] The advantage of the foregoing is that the translational look-aside buffer can be used to store both virtual addresses including the thread implicit bit for addresses related to a thread of a multi thread processor, or to store a virtual address which relates to a single thread processing system.

[0023] The foregoing apparatus for storing and utilizing translational look-aside buffer (TLB) entries which include a thread implicit (TI) bit carries out the process shown in Fig. 4. Referring now to Fig. 4, the process of reading data from

the translational look-aside buffer (TLB) begins in step 41. The process requires a determination of a thread ID associated with a particular virtual address being requested by the processor. The virtual address (VA) includes an effective address (EA) and a process identifier (PID). In a multi thread processor such as shown in Fig. 1, a thread ID will be available from the thread ID register 12.

[0024] The hardware engine 25 of the translational look-aside buffer (TLB) determines the VPN in step 43, representing the process identifier PID (0:7) and the first 22 bits of the effective address EA (0:21) requested by the process being executed in the CPU pipeline processor.

[0025] Once the VPN is known, the real address associated with the VPN comprising both higher order RPN (0:22) and lower order real address bits RPN (22:25) stored with the virtual page number (VPN) are determined in step 44. The determination is made in decision block 45 as to whether or not the thread implicit mode bit Tbit has been set. If the bit has been set, indicating that the address sought is particular to a specific thread ID, process step 46 replaces the lower order real address bits RPN (22:25) with the thread ID in step 46 by concatenating the thread ID with the higher order real address bits RPN (0:22). As a final step, the real address is concatenated in step 47 with the lower order bits of the effective address EA (22:31).

[0026] If decision block 45 determines that the thread implicit bit is not set to 1, representing a conventional look-aside buffer entry, then the translational look-aside buffer (TLB) contents are processed by using all of the real address data bits

concatenated with the effective address (EA) lower order data bits in step 47.

[0027] The foregoing description of the invention illustrates and describes the present invention. Additionally, the disclosure shows and describes only the preferred embodiments of the invention in the context of a translation look aside buffer (TLB) with increased translational capacity for multi-threaded computer processes, but, as mentioned above, it is to be understood that the invention is capable of use in various other combinations, modifications, and environments and is capable of changes or modifications within the scope of the inventive concept as expressed herein, commensurate with the above teachings and/or the skill or knowledge of the relevant art. The embodiments described hereinabove are further intended to explain best modes known of practicing the invention and to enable others skilled in the art to utilize the invention in such, or other, embodiments and with the various modifications required by the particular applications or uses of the invention. Accordingly, the description is not intended to limit the invention to the form or application disclosed herein. Also, it is intended that the appended claims be construed to include alternative embodiments.